DOCUMENT RESUME

ED 212 939 CG 015 733

AUTHOR Hoffman, Tony
TITLE On Being Sensitive and Specific About Child Abuse: Problems in Preventive Testing for Abusiveness in America.
SPONS AGENCY National Institutes of Health (DHEW), Bethesda, Md. Div. of Research Resources.
PUB DATE Apr 81
GRANT BRSG-SO-7-RR-07080
NOTE 18p.; Paper presented at the Annual Meeting of the Western Psychological Association (61st, Los Angeles, CA, April 9-12, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Child Abuse; Family Characteristics; Family Life; Parent Child Relationship; Parent Role; *Predictive Validity; Prevention; *Psychological Patterns; *Psychometrics; *Screening Tests; State of the Art Reviews; *Test Validity

ABSTRACT

Literature on psychological characteristics of abusive families suggests that psychometric assessments can be developed to screen parents for abusiveness. Advocates of primary prevention of child abuse have encouraged the development of such perinatal screening devices. However, social values define abusiveness, making its assessment especially prone to error. If a nearly perfect test existed with a sensitivity of .99 and a specificity of .99, half of the parents identified as abusive would not be abusive. Existing tests have done far worse, achieving at best a sensitivity of .77 and a specificity of .83. Although the low base rate of abuse prevents standard validity coefficients from accurately portraying the validity of tests for abusiveness, current tests appear to have little validity and high rates of predictive error. Given various levels of test precision, probabilities for correct and incorrect assessments of abusiveness can be determined; errors in virtually any test for abusiveness will be false positives, i.e., a detection of abusiveness when none exists. Considering the ethical and economical issues involved in testing the general population and intervening upon nonabusive families identified as abusive as well as the obstacles to gaining predictive validity in tests, an abandonment of testing efforts warrants serious consideration. (Author/NRB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED212939

ON BEING SENSITIVE AND SPECIFIC ABOUT CHILD ABUSE: PROBLEMS IN PREVENTIVE TESTING FOR ABUSIVENESS IN AMERICA

Tony Hoffman

Department of Psychology

University of Oregon

Eugene, Oregon 97403

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Tony Hoffman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the annual meeting of the
Western Psychological Association
in Los Angeles, April 9, 1981

This study was supported in part by BRSG Grant SO7RR07080, awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health.

This paper was inspired by studies of psychometric assessment with Lewis Goldberg. Programming by Ron Friedland aided in the computation of the conditional probabilities in this paper.

The help of Robin Mermelstein and Beverly Fagot in supporting this research is greatly appreciated. Special thanks go to Lewis Goldberg, Richard Littman, and Robin Mermelstein for their comments on earlier drafts of this paper. Errors and misrepresentations in this paper are the sole responsibility of the author.

CG 015733

On Being Sensitive and Specific About Child Abuse: Problems in Preventive Testing for Abusiveness in America

The forthcoming National Study of the Incidence and Severity of Child Abuse and Neglect (NSISCAN) conservatively estimates that one million children are maltreated in America (Note 1). Only 44 percent of these children are actually referred to child protective services (CPSs) and other investigatory agencies. Our growing concern for more accurate methods of predicting and preventing this maltreatment is certainly warranted. Unfortunately, this paper shows that existing predictive tests for maltreatment have very high error rates and little validity; it also shows that the tests must be extraordinarily sensitive and specific to have much validity in predicting abuse (cf. Schneider, Helfer & Hoffmeister, 1980).

Many children's advocates are concerned with the primary prevention of maltreatment (e.g., Miller, 1981), which involves the prediction and prevention of maltreatment prior to its actual occurrence. Assuming that tendencies toward maltreatment can be detected by assessing psychological characteristics, a number of research attempts to predict maltreatment have used psychometric assessment techniques (Schneider et al., 1980). The techniques are primarily designed as perinatal screening devices that can be efficiently used with mothers in hospital and home settings.[1]

The techniques detect "poor parenting" practices (Schneider et al., 1980), which include the neglect and abuse of children,[2,3] In this paper, child abuse and neglect by parents is called "abuse" and the propensity for these acts is called "abusiveness." The assessment devices for detecting abusiveness include psychometric testing, behavioral assessment and screening, clinical diagnosis; among others. This paper is concerned with all of these detection techniques, but refers to them as "tests" or "psychometric instruments" for abusiveness. Its comments regarding the testing for abusiveness can also be generalized to any testing of phenomena that have a low incidence in the general population (e.g., Rosen, 1954; Lykken, 1981; see also Wiggins, 1973).

Some types of physical violence toward children are acceptable (Gil, 1970; Viano, 1974) and frequently used in American families (Gelles, 1978; Steinmetz, 1977; Straus, Gelles & Steinmetz, 1980; see also Geis & Monahan, 1975). Abuse that would warrant referral to a CPS agency, however, has a relatively low

incidence. A high (but commonly used--Starr, 1979) incidence estimate is that, yearly, one percent of America's children suffer such abuse.[4] A one percent rate of parental abusiveness is assumed for the didactic purposes of this paper; this estimate is high according to many studies, and would constitute an enormous CPS caseload if fully detected.[5]

Detecting Abusiveness

The concept of abusiveness as a discrete form of psychopathology has been replaced (Gelles, 1973; Parke & Collmer, 1975; Spinetta & Rigler, 1972) by evidence suggesting that abusiveness is distributed along a continuum of family interaction patterns. "True" abuse is behavior that social values determine to be unacceptably abusive (Giovannoni & Becerra, 1979). This means that the behavior is judged to surpass a social criterion of abusiveness. If it is assumed that abusiveness is distributed along a continuum, then the task for tests of abusiveness is to discriminate parents whose abusiveness would be above a social criterion, and those who would be below this criterion.

The criterion for most tests of abusiveness is the determination that a parent's abusive behavior warrants a CPS referral.[6] This paper evaluates exemplary tests for their ability to predict such behavior.

Prediction Errors

A test for abusiveness must correctly predict actual abuse (true or valid positive cases of abusiveness) and correctly predict actual nonabuse (true or valid negative cases of abuse). If the test makes predictive errors, these will be of two types: false negatives, where truly abusive families are predicted to be nonabusive, and false positives, where truly nonabusive families are predicted to be abusive.

The scores resulting from testing are distributed along a continuum of predicted parental abusiveness. Parents can then be dichotomized as abusive or nonabusive using a cutoff at some point along the continuum, above which a family is deemed "abusive," and below, "nonabusive."[7]

Figure 1 shows the distributions of scores for a hypothetical predictor of abusiveness. The frequency of both groups' scores are assumed to be normally distributed. Child abusers (true positives) as a group score higher on the predictor than nonabusers (true negatives).[8] But there is some overlap in the distributions. Some nonabusers have the same scores as abusers; the test has some

degree of predictive error. The cutoff score is set so that any scorer above the cutoff is predicted to be abusive and any scorer below the cutoff is predicted to be nonabusive.[9] The predictor errs to the extent that abusers fall below the cutoff (false negatives) and to the extent that nonabusers fall above the cutoff (false positives).

Both types of error have serious social ramifications. A false negative error (a failure to detect true abusiveness) may cause serious danger to the life of a child who is being abused. A false positive error (a detection of abusiveness when none is present) may result in serious invasions upon the privacy of nonabusive families, and be a waste of valuable protective service time that could otherwise be spent on truly abusive families.

The art of test development is to minimize both false positive and false negative errors. The probability of the errors are linked, however, so that a prediction technique that minimizes one type of error will tend to increase the other. Increasing the power of a test by decreasing false negative errors lowers the cutoff (shift it to the left in Figure 1). But this automatically increases the frequency of false positive errors. Similarly, minimizing false positive errors maximizes false negative errors.

For example, a criterion predicting that abused children are those that have been seriously burned three or more times will have a relatively low number of false positive errors (few children are seriously burned three or more times by accident), and a relatively high number of false negative errors (many abused children that are subjected to burns are seriously burned but once or twice before their abusive parents realize they may end up in court). In an attempt to decrease these false negative errors, lowering the abuse criterion to two or more serious burns will increase false positive errors (more children are accidentally burned two times than three) and decrease false negative errors (there are fewer abused children who are burnt once than there are abused children who are both burnt once or twice). Such a criterial change may be attractive for a CPS agency seeking to prevent as much abuse as possible. The cost of this policy, however, would involve the increased burden of attending to more nonabusive families that would not have burnt their children any more (i.e., the increased number of false positive errors.)

Difficulties in Predicting Abuse

Assuming that a test for abusiveness would be administered to the general population, how difficult will it be to discriminate that one percent of abusive

families from the 95 percent that are not?

Test accuracy can be evaluated by analyzing the degree to which a test correctly detects "true" abusiveness and nonabusiveness (the abusiveness and nonabusiveness defined by society). The sensitivity of a test reflects its success at valid positive predictions, and is defined as the proportion of correctly detected abusive parents. The specificity of a test reflects its success at valid negative predictions, and is defined as the proportion of correctly detected nonabusive families. These two proportions can be used to evaluate the predictive validity of a test for abusiveness when the incidence of abuse is known or estimated.[10]

Assuming a .01 base rate of abuse, Table 1 gives the proportion of true positive cases of abuse that will be detected at varying degrees of test sensitivity and specificity (computed using Bayes theorem). Even an optimal screening instrument with .95 sensitivity (detecting 95 percent of all actual cases of abuse) and .95 specificity (detecting 95 percent of all true cases of nonabuse) would correctly diagnose only .16 of the cases predicted to be abusive. Eighty-four percent of the cases predicted to be abusive would actually be false positives.

This may seem counter-intuitive. But, considering the low incidence of abuse, even the slightest proportion of errors in detecting nonabusers creates a large portion of positive cases that are false. Thus, even a test with .99 specificity and .99 sensitivity will diagnose abusiveness incorrectly 50 percent of the time.

The best tests of abusiveness have done far worse. Since 1966, Helfer and his colleagues (Schneider, Hoffmeister & Helfer, 1976; Helfer, Schneider & Hoffmeister, Note 6) have been developing the Michigan Screening Profile of Parenting (the MSPP), a perinatal screening questionnaire (Helfer, Hoffmeister & Schneider, Note 6). Table 2 is a sample of item clusters used in the MSPP. After at least four revisions of the instrument (and its administration to over 500 mothers) Helfer, Schneider, and Hoffmeister (Note 6) reported that their best cluster of items had a sensitivity of .77 and a specificity of .83, when a criterion of having at least one CPS referral for abusiveness was used (cited in Starr, 1979).[11] Other screening devices have been developed (e.g., Disbrow, Doerr & Caulfield, 1977; Gaines, Sangrund, Green & Power, Note 7; Starr, Ceresnie & Steinlauf, Note 8) using discriminant function analysis, with similar results (Starr, 1979). Starr projected that an instrument with a sensitivity of .85 and a specificity of .83 could be developed (Starr, 1979).[12]

Assuming a .01 base rate of abuse, Helfer et al's questionnaire would correctly diagnose .04 of the parents it predicts to be abusive, and .96 would be false positives. Starr's projections would have .05 valid positives, and .95 false positives. Clearly, such instruments will do little to reduce CPS errors in predicting abusiveness; what is particularly upsetting about Table 1 is that an extraordinary amount of sensitivity and specificity are needed to reduce the proportion of false positives to even .5.

The Nature of the Detection Errors

Table 3 shows the proportion of detection errors that would occur in the total population at a range of sensitivities and specificities. At .99 specificity and .99 sensitivity, .01 of the population would be misclassified. Considering that .01 of the population is assumed to be abusive in the table, a large proportion of misclassifications must be anticipated even when almost perfect instruments are used. A test of .95 specificity and .95 sensitivity would misclassify .051 of the population. Helfer et al's MSPP would misclassify .172 of the population and Starr's (1979) projections would misclassify .175 of the population (both over 17 times the actual rate of abuse).

By far the largest proportion of the detection errors using any valid instrument will be false positive errors. Table 4 shows the proportion of errors that would be false positives, assuming a .01 base rate of abuse. The great majority of predictive errors are false positive errors. Only when specificity is greater than .99 and sensitivity is less than .90 is the proportion of errors that are false positives less than .90.

Table 4 shows that errors in virtually any test for abusiveness will be false positives; Table 1 shows that the great proportion of positive predictions will be false. With the exception of nearly perfect (and currently nonexistent) valid tests for abusiveness, it must be anticipated that the great majority of parents predicted to be abusive are probably not.

Conclusion

In the enterprise of psychometric screening, a test whose sensitivity and specificity exceed .90 is hard to find. It is truly extraordinary to have a test with these values exceeding .95.

A truly extraordinary amount of sensitivity and specificity is needed if a test for abusiveness is to have a satisfactory amount of validity. The test

would have to be much more sensitive and specific than similar psychometric tests (Wiggins, 1973) and would have to employ discriminative techniques yet to be developed in the art of psychometrics (Nunnally, 1978; Wiggins, 1973).

Due to the low base rate of abuse, standard validity coefficients cannot accurately portray the validity of a test for abusiveness. It can nevertheless be concluded that current tests of abusiveness have little validity and high rates of predictive error. The prospects for developing valid and useful tests for abusiveness are poor (cf. Schneider et al., 1980).

Considering that such tests are needed for screening in primary prevention efforts, the primary prevention of child abuse may be a crusade doomed for failure (cf. Zigler, 1979, 1980). The large proportion of false positive families that would be screened for CPS investigation by these tests raises the serious ethical issue of intervening upon a large number of nonabusive families (Gil, 1976a, 1976b). Screening for primary prevention also raises the ethical issue of intervening in abusive families prior to the occurrence of any actual abuse.

Given the high false positive error rates that optimal tests of abusiveness may carry, any primary prevention effort using the tests as screening devices will be expending a large amount of protective service efforts on nonabusive families. Given the limited funds of many CPSs, such an effort might not adequately address the problems of the relatively few abusive families that would be included in the program. If the primary prevention effort were to drain funds from secondary prevention programs, truly abusive families may not have their problems addressed in either primary or secondary programs.

The value of a primary prevention effort, given its inability to concentrate on abusive families, may actually be detrimental to abuse prevention efforts if secondary programs are weakened in the effort. Considering that secondary programs concentrate on abusers and that primary prevention programs cannot (given their high error rate), it may be more valuable and ethical to emphasize secondary prevention efforts than to experience the predictive failures of a primary prevention effort (cf. Gil, 1976b; Zigler, 1979, 1980). An abandonment of the testing effort warrants serious consideration.

Footnotes

[1]It is important to note that, even though fathers are responsible for more than half of all reported parental abuse (Note 2), fathers are rarely screened in such research.

[2]Sexual abuse is not considered in this paper. This is not meant to imply that sexual abuse is of no concern, but that sexual abuse is unlikely to be detected with tests for poor parenting. One-third of sexual abuse is perpetrated by nonparents (Note 1) and often occurs in the absence of overt signs of physical abuse, neglect, or behavior dysfunction on the part of adults (Kempe, 1980).

[3]The traditional categorization of maltreatment into physical abuse, neglect, and sexual abuse are admittedly indistinct for assessment purposes. The current attempt at assessments of poor parenting practices taps at a more general feature of family problems associated with child maltreatment. This attempt may be criticized for the culture-bound vagueness of criteria for poor parenting.

Community members, protective service workers, and professionals recognize at least five distinct categories of child maltreatment (Giovannoni & Becerra, 1979). More specific attempts at detecting maltreatment tendencies might profit from discriminating between the categories of physical abuse, sexual abuse, failure-to-provide, supervision problems, and parental substance abuse.

[4]Community estimates of the rate of abuse range no higher than one percent (see, e.g., Light, 1973; Note 3). Other accounts have, with one exception (Fontana, 1973) suggested incidence rates lower than one percent (Gelles, 1978; see, e.g., Cohen & Sussman, 1975).

The one percent figure is also high as a base rate of parental abusiveness. Although it reflects an estimate of the rate of abuse, it does not compensate for the fact that many of these children were not abused by parents. Also, many of these children have two parents, and it is possible that only one of these parents was abusive. Thus, the one percent estimate of parental abusiveness is high because there are fewer abusive parents than abused children. It is set at this level in this paper for the sake of didactic simplicity. The conclusions and considerations in this paper can be more strongly emphasized if the rate of parental abusiveness is lower than the assumed base rate of one percent.

The NSISCAN (Note 1) found that roughly one percent of America's children are recognized by CPSs to be abused each year, using criteria established by the National Center on Child Abuse and Neglect (Note 4). However, the extent of undetected cases of abuse warrants much higher projections of the extent of abusiveness (but see footnote 5).

The primary concern in this paper is with demonstrating problems in detecting abusiveness at a high, but realistic, estimate of the rate of abuse that might be attended to in CPS caseloads. Presumably, we should seek to detect the most abusive population that CPSs can accomodate. One percent of America's family population is an enormous caseload; even if the projected rate of abuse is higher than one percent, detecting the one percent that might receive CPS attention is at issue.

[5]The NSISCAN (Note 1) actually projects a base rate of abuse as high as 5 percent, based on estimates of reporting frequencies to CPSs. Unfortunately,

the reporting frequencies of the many referral sources that abused children may be contacted by are not summed in the projection. It may therefore be a high and liberal estimate of the extent of abuse in America.

[6]In turn, as social control institutions, the courts and CPSs use criteria for abusiveness that reflect dominant community values. These criteria are vaguely defined, despite high professional and community agreement on acts that constitute abuse (Giovannoni & Becerra, 1979; Lena & Warkov, 1978; Note 5).

Published accounts of tests of abusiveness use criteria for abusiveness that are often unspecified. Unfortunately, a number of studies (e.g., Gray et al., 1977) do not specifically state if the referrals were substantiated cases of abuse; typically, over half of all referrals to a CPS agency are not substantiated. Other criteria, such as "poor parenting" are also used (e.g., Elmer, 1977), with vague criteria.

[7]This is not denying the continuum of family relations (no family is completely abusive or completely nonabusive). The considerations here are with society's assessment of abusiveness, where only two choices are made (and that are therefore dichotomous): either abuse is present in a family, or it is not. The decision, even if it is just for the purposes of protective screening, is dichotomous: either the family is to be investigated for abusiveness, or it is not.

[8]Note that the distributions of abusers and nonabusers are from equal sized samples; representative sampling would result in a distribution of nonabusers that is almost 99 times larger than that of abusers.

[9]The cutoff score need not be set where the frequencies of the two distributions on a predictor score are equal, as in Figure 1. It is often set at a point where the total number of predictive errors are minimized, but may be set at a point where one type of error is minimized more than the other.

[10]This approach to test validity has been used for some time, and many of the analyses and validity considerations are derived from the work of Taylor and Russell (1939). Table 1 is a variation of what are called "Taylor-Russell" tables.

[11]The Helfer, Schneider, and Hoffmeister (Note 6) manuscript is unpublished and was not made available to the author upon request; these sensitivity and specificity values are reported from the secondary source of Starr (1979). Schneider et al. (1980) report, however, that the MSPP has "approximately an 80% sensitivity and specificity in predicting known abuse . . ." (pp. 424-425).

[12]Other assessment techniques have been developed. The MSPP (Helfer et al., 1978) and Starr's projections were selected for critique in this paper. The former is the most widely used screening device for abusiveness (and apparently has been studied under a number of replications), and the latter has the highest realistic sensitivity and specificity projections known to the author.

Reference Notes

1. Recognition and reporting of child maltreatment: Findings from the National Study of the Incidence and Severity of Child Abuse and Neglect. Advance report prepared for the National Center on Child Abuse and Neglect; Children's Bureau; Administration for Children, Youth, and Families; U.S. Department of Health and Human Services; Washington, D.C. 20201, December, 1980.

2. National analysis of official child abuse and neglect reporting, 1978. National Center on Child Abuse and Neglect, U.S. Department of Health, Education and Welfare, 1979. Children's Bureau; Administration for Children, Youth, and Families; U.S. Department of Health and Human Services; Washington, D.C. 20201, 1980.

3. Nagi, S. Z. The structure and performance of programs on child abuse and neglect: Interim report: Child abuse and neglect programs--a national interview. Prepared for the Office of Child Development; U.S. Department of Health, Education and Welfare, Washington, D.C. 20201, March 1975.

4. Technical report number 2: Operational definition of child maltreatment. National study of the incidence and severity of child abuse and neglect. Submitted to the National Center on Child ABuse and Neglect; Children's Bureau; Administration for Children, Youth, and Families; U.S. Department of Health and Human Services; Washington, D.C. 20201, 1980.

5. Gelles, R. J. Problems in defining and labeling child abuse. Paper presented to the Study Group on Problems in the Prediction of Child Abuse and Neglect. Wilmington, Del., June 1977.

6. Helfer, R. E., Schneider, C. J., & Hoffmeister, J. K. Report on research using the Michigan Screening Profile of Parenting (MSPP). Unpublished manuscript. East Lansing: Michigan State University, May 1978. Cited in Starr (1979).

7. Gaines, R., Sandgrund, A., Green, A. H., & Power, E. Etiological factors in child maltreatment: A multivariate study of abusing, neglecting, and normal parents. Brooklyn, N.Y.: Downstate Medical Center, 1977.

8. Starr, R. H., Jr., Ceresnie, S. J., & Steinlauf, B. Social and psychological characteristics of abusive mothers. Paper presented at the meeting of the Eastern Psychological Association, Washington, D.C., May 1978.

References

Cohen, S., & Sussman, A. The incidence of child abuse in the U.S. Child Welfare, 1975, 54, 432-443.

Disbrow, M. A., Doerr, H., & Caulfield, C. Measures to predict child abuse. Paper presented at the meeting of the Society for Research in Child Development, New Orleans, March 1977.

Elmer, E. Fragile families, troubled children: The aftermath of infant trauma. Pittsburgh: University of Pittsburgh Press, 1977.

Gelles, R. J. Child abuse as psychopathology: A sociological critique and reformulation. American Journal of Orthopsychiatry, 1973, 43(4), 611-621.

Gelles, R. J. Violence towards children in the United States. American Journal of Orthopsychiatry, 1978, 48, 580-592.

Geis, G., & Monahan, J. The social ecology of violence. In T. Lickona (Ed.), Man and morality. New York: Holt, Rinehart, & Winston, 1975.

Gil, D. G. Violence against children: Physical child abuse in the United States. Cambridge, Mass.: Harvard University Press, 1970.

Gil, D. G. Primary prevention of child abuse: A philosophical and political issue. Psychiatric Opinion, 1976, 13(2), 30-34. (a)

Gil, D. G. The challenge of social equality. Cambridge, Mass.: Schenckman Press, 1976. (b)

Giovannoni, J. M., & Becerra, R. M. Defining child abuse. New York: The Free Press, 1979.

Gray, J. O., Cutler, C. A., Dean, J. G., & Kempe, C. H. Prediction and prevention of child abuse and neglect. Child Abuse and Neglect, 1977, 1, 45-58.

Helfer, R. E., Hoffmeister, J. K., & Schneider, C. J. Michigan Screening Profile of Parenting. Boulder, Colorado: Test Analysis and Development Corp., 1978.

Kempe, C. H. Incest and other forms of sexual abuse. In C. H. Kempe & R. E. Helfer (Eds.), The battered child. Chicago: The University of Chicago Press, 1980.

Lena, H. F., & Warkov, S. Occupation perceptions of the causes and consequences of child abuse/neglect. Medical Anthropology, 1978, 2(1), 1-28.

Light, R. J. Abused and neglected children in America: A study of alternative policies. Harvard Educational Review, 1973, 43(4), 556-598.

Lykken, D. A tremor in the blood: Use and abuses of the lie detector. San Francisco: McGraw-Hill, 1981.

Miller, C. C. Primary prevention of child mistreatment: Meeting a national need. Child Welfare, 1981, 60(1), 11-23.

Nunnally, J. C. Psychometric theory (2nd ed.). New York: McGraw-Hill, 1978.

Parke, R. D., & Collmer, C. W. Child abuse: An interdisciplinary analysis. In E. V. Hetherington (Ed.), Review of child development research (Vol. 5). London: University of Chicago Press, 1975.

Rosen, A. Detection of suicidal patients: An example of some limitations in infrequent events. Journal of Consulting Psychology, 1954, 18, 397-403.

Schneider, C., Helfer, R. E., & Hoffmeister, J. K. Screening for potential abuse: A review. In C. H. Kempe & R. E. Helfer (Eds.), The battered child (3rd ed.). Chicago: University of Chicago Press, 1980.

Schneider, C., Hoffmeister, J. K., & Helfer, R. E. A predictive screening questionnaire for potential problems in mother-child interaction. In R. E. Helfer & C. H. Kempe (Eds.), Child abuse and neglect: The family and the the community. Cambridge, Mass.: Ballinger, 1976.

Spinetta, J. J., & Rigler, D. The child abusing parent: A psychological review. Psychological Bulletin, 1972, 77, 296-304.

Starr, R. H., Jr. Child abuse. American Psychologist, 1979, 34(10), 872-878.

Steinmetz, S. K. The use of force for resolving family conflict: The training ground for abuse. The Family Coordinator, 1977, 26, 19-26.

Straus, M. A., Gelles, R. J., & Steinmetz, S. K. Behind closed doors: Violence in the American family. Garden City, N.Y.: Anchor Press, 1980.

Taylor, H. C., & Russell, J. T. The relationship of validity coefficients to the practical effectiveness of tests in selection. Journal of Applied Psychology, 1939, 23, 565-578.

Wiggins, J. S. Personality and prediction: Principles of personality assessment. Reading, Mass.: Addison-Wesley, 1973.

Zigler, E. Controlling child abuse in America: An effort doomed to failure? In R. Bourne & E. H. Newberger (Eds.), Critical perspectives on child abuse. Lexington, Mass.: Lexington Books, 1979.

Zigler, E. Controlling child abuse: Do we have the knowledge and/or the will? In G. Gerbner, C. J. Ross, & E. Zigler (Eds.), Child abuse: An agenda for action. New York: Oxford University Press, 1980.

Errata:
Fontana, V. Somewhere a child is crying. New York: Macmillan, 1973.

Viano, E. C. Attitudes towards child abuse among American professionals. Paper presented at the biennial meeting of the International Society for Research on Aggression, Toronto, Canada, 1974.

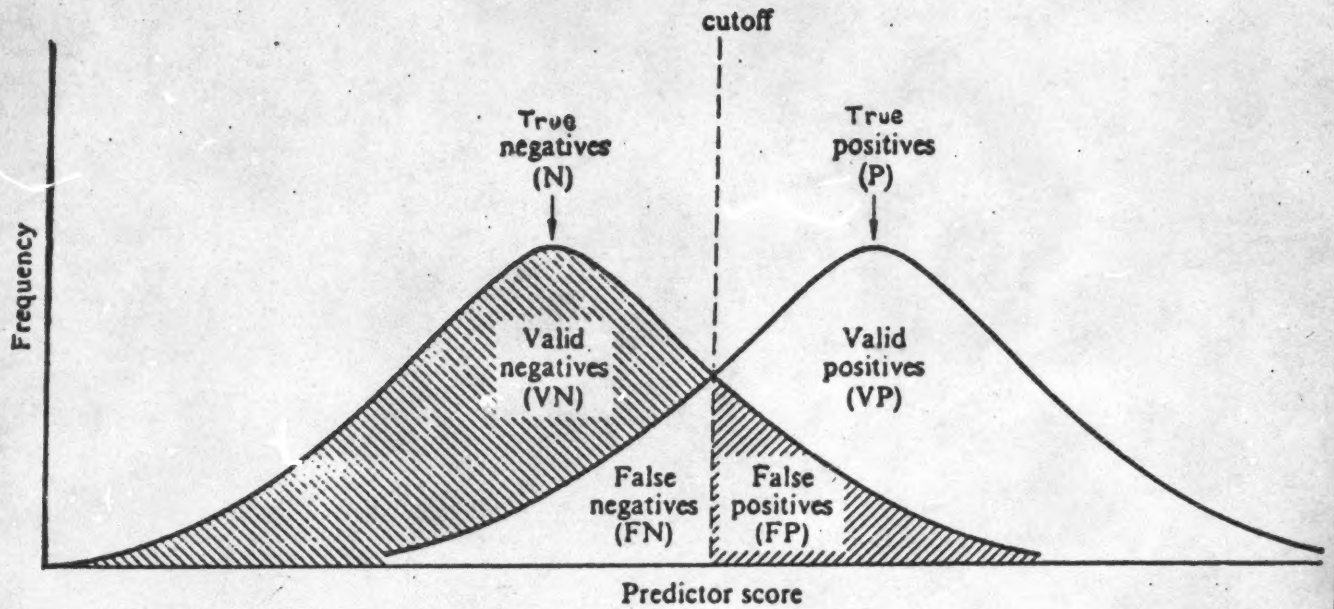

Figure 1. Distributions of scores on a predictor of abusiveness for equal-sized populations of abusers (true positives) and nonabusers (true negatives). These hypothetical distributions are assumed to be normally distributed. True abusers that fall above the cutoff are valid positive detections of abusiveness, and true nonabusers that fall below the cutoff are valid negative detections. Some true abusers fall below the cutoff and are predicted to be nonabusive (false negatives); some true nonabusers fall above the cutoff and are predicted to be abusive (false positives). See text for details.

TABLE 1

PROPORTION OF PARENTS PREDICTED TO BE ABUSIVE BY A PSYCHOMETRIC INSTRUMENT WHO ARE TRULY ABUSIVE[a]

	Specificity[c]							
Sensitivity[b]	.50	.60	.70	.80	.90	.95	.99	1.00
.50	.010	.012	.016	.025	.048	.092	.336	1.000
.60	.012	.015	.020	.029	.057	.108	.377	1.000
.70	.014	.017	.023	.034	.066	.124	.414	1.000
.80	.016	.020	.026	.039	.075	.139	.447	1.000
.90	.018	.022	.029	.043	.083	.154	.476	1.000
.95	.019	.023	.031	.046	.088	.161	.490	1.000
.99	.020	.024	.032	.048	.091	.167	.500	1.000
1.00	.020	.025	.032	.048	.092	.168	.502	1.000

[a] assuming a .01 base rate of parental abuse. This is the proportion of valid positives among predicted positive cases of abuse. Computed using Bates' Theorem; the different sensitivies[b] and specificities[c] are those of the hypothetical test involved.

[b] the proportion of truly abusive parents correctly predicted to be abusive by the psychometric instrument.

[c] the proportion of truly nonabusive parents correctly predicted to be nonabusive by the psychometric instrument.

TABLE 2

SAMPLE ITEM CLUSTERS FROM THE MICHIGAN SCREENING PROFILE OF PARENTING

CLUSTER I: Problems with Mother (PWM)
My mother and I have always gotten along well.
The main thing I remember from childhood is the love and warm affection my parents showed me.

CLUSTER II: I'm No Damn Good (IM)
No one has ever really listened to me.
I am always being criticized by other people.
I have never really felt loved.

CLUSTER III: Isolation (ISO)
I am close to other people.
I am always good to other people.
I am very well liked by everyone.

CLUSTER IV: Child Expectations--A(CEA)
Children should know even before the age of 2 years what parents want them to do.
Children need to be taught before the age of 2 to respect and obey their parents.

CLUSTER V: Child Expectations--B(CEB)
I never become angry with my child(ren).
My child(ren) is (are) always good.
Most children should walk well by 9 or 10 months of age.

CLUSTER VI: Depression/Crisis (DEP)
I am afraid of many things.
Often when my baby cries, I don't know what to do about it.
Sometimes I just feel like running away.

Gut Score (GUT)
Children should know even before the age of 2 years what parents want them to do.
As a child I often felt no one paid much attention to what I wanted or needed.
Often when my baby cries, I don't know what to do about it.
When I was a child my parents used severe physical punishment on me or one of my siblings.

Discriminant Items (DIS)
When I was a child my parents used severe physical punishment on me or one of my siblings.
At least one of my parents didn't really listen to me or understand my feelings.
It is extremely important for me to have my children behave well even when they are infants.

From C. Schneider, J. Hoffmeister, & R. E. Helfer. A predictive screening questionnaire for potential problems in mother-child interaction. In R. E. Helfer & C. H. Kempe (Eds.), Child abuse and neglect: The family and the community. Cambridge, Mass.: Ballinger Publishing Co., 1976, pp. 396-397.

TABLE 3

PROPORTION OF DETECTION ERRORS IN TOTAL POPULATION AT HIGH SENSITIVITIES AND SPECIFICITIES[a]

	Specificity[c]							
	.50	.60	.70	.80	.90	.95	.99	1.000
Sensitivity[b]								
.50	.500	.401	.302	.203	.104	.055	.015	.005
.60	.499	.400	.301	.202	.103	.054	.014	.004
.70	.498	.399	.300	.201	.102	.053	.013	.003
.80	.497	.398	.299	.200	.101	.052	.012	.002
.90	.496	.397	.298	.199	.100	.051	.011	.001
.95	.496	.397	.298	.199	.100	.051	.011	.001
.99	.495	.396	.297	.198	.099	.050	.010	.000
1.00	.495	.396	.297	.198	.099	.050	.010	0.000

[a] the sum of the proportion of false positive and false negative errors, or the sum of the proportion of American parents that were detected to be abusive when they actually were not, plus the proportion of American parents that were detected to be nonabusive when they actually were abusive; using a psychometric instrument for abusiveness on a population with a .01 base rate of parental abuse.

[b] proportion of abusive parents that are predicted to be abusive.

[c] proportion of nonabusive parents that are predicted to be nonabusive.

TABLE 4

PROPORTION OF DETECTION ERRORS THAT ARE FALSE POSITIVES AT HIGH SENSITIVITIES AND SPECIFICITIES[a]

Sensitivity[b]	Specificity[c] .50	.60	.70	.80	.90	.95	.99	1.000
.50	.990	.986	.983	.975	.952	.908	.664	0.000
.60	.992	.990	.987	.980	.961	.925	.712	0.000
.70	.994	.992	.990	.985	.971	.943	.767	0.000
.80	.996	.995	.993	.990	.980	.961	.832	0.000
.90	.998	.997	.996	.995	.990	.980	.908	0.000
.95	.999	.999	.998	.997	.995	.990	.952	0.000
.99	.999	.999	.999	.999	.999	.998	.990	0.000
1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000	-----

[a] the proportion of false positives in the total errors, assuming a .01 base rate of parental abuse. See note a, Table 3.

[b] the proportion of abusive parents that are predicted to be abusive.

[c] the proportion of nonabusive parents that are predicted to be nonabusive.